

# THE INTERLANGUAGE SPEECH INTELLIGIBILITY BENEFIT FOR EXPERIENCED NONNATIVE LISTENERS: PERCEPTION OF ENGLISH LEXICAL STRESS PRODUCED BY KOREAN NATIVE SPEAKERS

*Jieun Song*

Seoul National University, Korea

dwflower1120@hotmail.com

## ABSTRACT

This study investigates the intelligibility of English lexical stress produced by Korean native speakers for English native listeners and Korean native listeners. Listeners had to listen to English disyllabic words where the position of stress determines whether the word is a noun or a verb and identify intended stress patterns.

The results reveal that the 'Interlanguage Speech Intelligibility Benefit' emerges when high-proficiency Korean listeners are involved. These experienced Korean listeners identified stress patterns produced by low proficiency Korean talkers significantly better than native English listeners. Acoustic analyses suggest that the native English listeners had difficulty understanding words where stress was marked only by F0 and intensity but without any appropriate reduction of vowel quality in unstressed syllables or difference in duration. The Korean high proficiency listeners, however, were very accurate at judging such stress patterns as intended as well as those produced by native English talkers. This will be discussed in relation to phonological systems of second language learners.

**Keywords:** intelligibility, English lexical stress, the 'interlanguage speech intelligibility benefit', second language perception

## 1. INTRODUCTION

There has been a great deal of research investigating factors affecting the intelligibility of nonnative speech. Recently, the language backgrounds and proficiency of talkers and listeners were found to affect the intelligibility of foreign accented speech. Specifically, the 'interlanguage speech intelligibility benefit (ISIB)' [1] refers to the situation where nonnative listeners, usually those who share their L1 with talkers have an advantage in understanding the nonnative speech. It was also shown that

intelligibility depends on the amount of experience L2 listeners and talkers have in the target language. For example, a few studies provided evidence that such benefit held only for high-proficiency talkers. [1] Regarding listeners' proficiency, it was found that the ISIB emerged when low proficiency nonnative listeners understood nonnative talkers who shared the L1 with them. [1, 5, 9] It was suggested that the intelligibility benefit is based on shared phonological knowledge between listeners and talkers. [2] However, a recent study indicated that French experienced listeners who were fairly fluent in English performed better than native English listeners in understanding English with French accented prosody when the noise level increased. [8]

The first goal of this study is to quantify the intelligibility of English produced by Korean native speakers, especially their production of English lexical stress, through a forced-choice identification test using minimal pairs like *CONtract*(noun) and *conTRACT*(verb). In English a stressed syllable is usually produced with higher pitch, louder intensity and longer duration than an unstressed syllable. In addition, most unstressed vowels are reduced to a schwa-like quality. Lee, et al. [6] showed that Korean early and late bilingual speakers of English used F0 in a native-like way to differentiate stressed and unstressed syllables, while they showed significantly less difference in duration and vowel quality between them than native English speakers. They explained this on the basis of the feature hypothesis which suggests that the L2 phonetic features not used in the L1, e.g., intensity, duration and vowel quality in this particular case, are more difficult to acquire than those that are used in the L1, i.e., F0.

Also, in order to shed light on the ISIB patterns which were found conflicting in previous studies, talker-listener interactions associated with the native language and proficiency of talkers and listeners will be explored. Furthermore, in an

attempt to understand why such interactions arise for certain groups of talkers or listeners, this study will analyze acoustic characteristics of English lexical stress produced by Korean native speakers.

## 2. METHODS

### 2.1. Talkers

Two native English speakers from North America (two males) and seven native Korean speakers (two males and five females) were recorded reading the materials used in the identification test. All of the Korean talkers had never lived in countries other than Korea and learned English mostly in school.

### 2.2. Materials

The talkers were recorded reading seven minimal pairs of English words used in [10]. Each minimal pair consists of a noun and a verb that differ only in their stress patterns. Stress falls on the first syllable when the word is a noun and on the second syllable when it is a verb. All Korean participants reported that they had learned about English lexical stress in school and most of them except for a few low proficiency talkers were aware that some English words are differentiated only by their stress patterns. All talkers were only told that the words in question should be read following the stress rules explained above and instructed to read context sentences used in [10] first so that they could remind themselves how those words should be pronounced. After that, they read each word three times in isolation. All productions were recorded with Sound Devices 722 at the sample rate of 44.1kHz(16bit) in a sound-attenuated room or quiet classrooms.

**Table 1:** Stimuli.

Words(noun/verb)
contract , desert, object, permit, rebel record, record

### 2.3. Listeners

Fifteen native English listeners from North America (twelve males and three females) and thirty Korean native listeners (thirteen males and seventeen females) participated in the forced-choice identification test. The Korean learners of English were undergraduate or graduate students and had not lived in any other country but Korea except for a few who had spent less than one year in English speaking countries at university. All listeners were first recorded reading the same materials used in the identification test so that their

proficiency in English phonology could also be evaluated.

### 2.4. Identification task

For every Korean native speaker, two tokens were selected for each word except for one speaker who occasionally exaggerated stress on purpose. For every English native speaker, only one token for each word was used for the test. A total of 220 tokens were selected for the identification test. They were played through headphones on a laptop computer in a random order. The listeners had to choose whether the word they heard was noun or a verb by clicking either the noun button or the verb button on the screen. The test was run in Praat5216 using an ExperimentMFC file. Along with the identification task, the listeners evaluated the degree of foreign accent of each stimulus on a 9-point scale.

### 2.5. Accentedness judgment task

The Korean listeners' proficiency in English phonology was evaluated as well in a following accentedness judgment task. Five phonetically trained researchers including the author and five native English speakers judged the degree of foreign accent of the 30 Korean listeners on the same 9-point scale. The evaluation was made on three context sentences produced by each listener out of five context sentences selected for this task.

## 3. RESULTS

### 3.1. Proficiency of talkers and listeners

The Korean talkers and listeners were classified as either high proficiency (HP) or low proficiency(LP) talkers/listeners according to their accentedness scores. In order to differentiate the proficiency levels of listeners more clearly, only six listeners who received the highest scores and six listeners who received the lowest scores were classified as HP listeners and LP listeners, respectively.

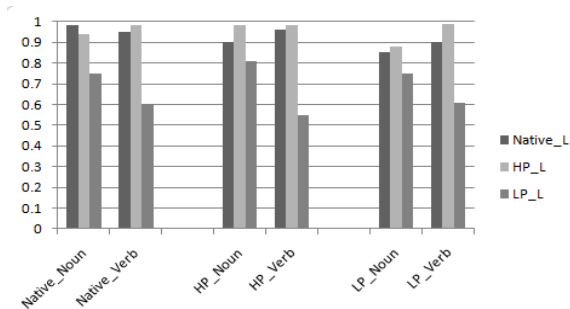
### 3.2. Word identification

Logistic regression analysis was carried out to predict correct identification of stress using proficiency of talkers and listeners, stress, the interaction between talker proficiency and listener proficiency, the interaction between listener proficiency and stress, the interaction between talker proficiency and stress, and the interaction

between talker proficiency, listener proficiency and stress as predictors. ( $\chi^2(11)=444.548$ ,  $df=11$ ,  $p<.001$ , Nagelkerke  $R^2=.139$ ) It was found that 'HP listener' produces the best fit of the data ( $b=3.126$  Wald  $\chi^2(1) = 69.289$ ,  $p<.001$ ), which indicates that HP listeners performed the best whether they listened to native English talkers or native Korean talkers. Next to HP listeners were native English (NE) listeners ( $b=1.099$ , Wald  $\chi^2(1) = 52.297$ ,  $p<.001$ ), which suggests that NE listeners were also relatively accurate at identifying stress. The interaction between NE listeners and NE talkers ( $b=1.245$ , Wald  $\chi^2(1)=9.618$ ,  $p<.05$ ) and the interaction between NE listeners and HP talkers ( $b=0.953$ , Wald  $\chi^2(1)=13.380$ ,  $p<.001$ ) were found significant as well. A MANOVA analysis was run in order to see if the ISIB occurred for HP listeners. HP listeners were significantly more accurate at understanding LP talkers than NE listeners. ( $F(1)=42.25$ ,  $p<.001$ )

Stress factor yielded significant results as well. First, 'Noun' was found to produce a significant fit of the data ( $b=0.634$ , Wald  $\chi^2(1)=12.898$ ,  $p<.001$ ). However, the interaction between NE listeners and noun ( $b= -0.575$ , Wald  $\chi^2(1)=6.288$ ,  $p<.05$ ) and the interaction between HP listeners and noun ( $b= -2.311$ , Wald  $\chi^2(1)=28.014$ ,  $p<.001$ ) were also significant. The negative factor weights imply that NE listeners and HP listeners failed to identify nouns to a greater extent than verbs.

**Figure 1:** identification rates of nouns/verbs by talker groups and listener groups.



**Table 2:** identification rates of nouns/verbs by talker groups and listener groups.

	NE Listener	HP Listener	LP Listener
	NOUN	NOUN	NOUN
NE Talker	0.98	0.94	0.75
HP Talker	0.9	0.98	0.81
LP Talker	0.85	0.88	0.75
	VERB	VERB	VERB
NE Talker	0.95	0.98	0.6
HP Talker	0.96	0.98	0.55
LP Talker	0.9	0.99	0.61

### 3.3. Acoustic analysis

In order to help illuminate why HP listeners identified Korean talkers, usually LP talkers significantly better than NE listeners, two groups of stimuli were acoustically compared. First, a set of fifteen tokens with which native listeners received the lowest identification rates (30-60%), but Korean HP listeners received 100% identification rates were selected, which will be called 'ISIB tokens'. These tokens were compared with tokens produced by native English speakers.

An F0 ratio was obtained from dividing average F0 of an unstressed syllable by that of a stressed syllable of each word. An independent T-test found no significant difference in mean F0 ratios between the two groups. This means that Korean talkers who produced the ISIB tokens were using F0 to differentiate stressed and unstressed syllables in a native-like way. In case of intensity ratios obtained the same way as F0 ratios, an independent T-test showed that the difference between the two groups was significant. ( $t(43)=2.848$ ,  $p<.05$ ) The ISIB-tokens did not reach the native-like intensity ratios.

Duration ratios obtained the same way, were compared for each word type since duration of a syllable partly depends on its syllable structure and segments that it consists of. For vowel quality, first formant and second formant of unstressed vowels which are expected to be reduced to a schwa-like quality were measured. They were compared for each group of words which were classified according to their underlying vowel quality. As shown in Table 4 and Table 5, ISIB tokens were quite different from native English speakers' tokens both in duration and in vowel quality. Mean duration ratios of the ISIB tokens did not reach those of native speakers. Unstressed syllables were produced with even longer duration than stressed syllables in many of the ISIB-tokens. While native English speakers reduced the unstressed vowels to schwa, the F1 and F2 values of the ISIB tokens were far from schwa. It seems that the Korean speakers retained the underlying quality of the vowels, which suggests that they had not learned to reduce unstressed vowels yet.

**Table 3:** mean F0 ratios and mean intensity ratios of (a) – ISIB-tokens (b)- native English speakers' tokens.

Acoustic cues	(a)	(b)
F0 ratio	.8539	.8406
Intensity ratio	<b>.9881</b>	<b>.9570</b>

**Table 4:** mean duration ratios of (a) – ISIB-tokens (b)- native English speakers' tokens.

word	(a)	(b)
desert(v.)	0.56	0.42
object(n.)	<b>1.9</b>	<b>1.05</b>
permit(n.)	0.82	0.96
permit(v.)	<b>1.05</b>	<b>0.64</b>
rebel(v.)	0.44	0.26
record(n.)	<b>3.48</b>	<b>1.76</b>
subject(v.)	<b>1.1</b>	<b>0.83</b>

**Table 5:** mean F1 and F2 values of (a) – ISIB-tokens (b)- native English speakers' tokens.

word	(a)	(b)
	<b>F1</b>	<b>F1</b>
desert, rebel(v.)	<b>559.71</b>	<b>358.54</b>
record(n.)	567	436.8
subject(v.)	<b>677.33</b>	<b>383.5</b>
	<b>F2</b>	<b>F2</b>
desert, rebel(v.)	<b>1935.86</b>	<b>1608.91</b>
record(n.)	<b>1158</b>	<b>1509.4</b>
subject(v.)	1569	1480.38

#### 4. DISCUSSION

This study examined the intelligibility of English lexical stress produced by Korean speakers of English. One of the major findings is that nonnative listeners who are relatively experienced in the target language performed significantly well in understanding both nonnative talkers who shared the L1 with themselves and native talkers. The interlanguage speech intelligibility benefit emerged especially when the HP listeners listened to LP talkers. The acoustic analyses showed that the HP listeners had no difficulty at all identifying the words where stress was signaled mostly by F0 without duration difference or vowel reduction. When adopting the view that nonnative speech intelligibility partly depends on how similar talkers' and listeners' phonological systems are to each other, it is assumed that the Korean HP listeners and the Korean LP talkers of this study share phonological representations of English lexical stress. In their representations of English, stressed syllables would be of higher pitch than unstressed syllables. However, given that the HP listeners were also as accurate at identifying native English speakers' production as native English listeners and they were in their later stage of acquiring English phonology, another possibility can be taken into account. That is, they might have acquired English lexical stress almost as much as native English speakers, but their experience in Korean speakers' stress patterns can allow them to make use of such knowledge when understanding

Korean accented English. Similarly, it was suggested in [8] that experienced L2 listeners can exploit their multilingual experience to have more flexible speech processing.

With respect to stress patterns, there was a tendency that NE listeners and HP listeners performed better in identifying verbs than nouns. In English, stress shifted from the left to the right, that is, SW pronounced as WS suffers in intelligibility to a greater degree than stress shifted from the right to the left. A leftward shift of stress is even a natural phenomenon in certain situations, especially for contrast.[3, 4] The tendency observed in this study might reflect that those two groups of listeners have a narrower range of representations about SW words compared to WS words which can often be pronounced as SW, which leads to more strict judgment about nonstandard pronunciation of nouns than verbs.

#### 5. REFERENCES

- [1] Bent, T., Bradlow, A. 2003. The interlanguage speech intelligibility benefit. *J. Acoust. Soc. Am.* 114(3), 1600-1610.
- [2] Bent, T., Bradlow, A., Smith, B. 2008. Production and perception of temporal patterns in native and non-native speech, *Phonetica* 65, 131-147.
- [3] Cutler, A., Clifton, C. 1984. The use of prosodic information in word recognition. In Bouma, H., Bouwhuis, D.G. (eds.), *Attention and Performance*. Hillsdale, NJ: Lawrence Erlbaum, 183-196.
- [4] Field, J. 2005. Intelligibility and the listener: The role of lexical stress, *TESOL QUARTERLY*. 39(3), 399-423.
- [5] Hayes-Harb, R., Smith, B.L., Bent, T., Bradlow, A. 2008. The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics* 36, 664-679.
- [6] Lee, B., Guion, S., Harada, T. 2006. Acoustic analysis of the production of Japanese bilinguals. *SSLA* 28, 487-513.
- [7] McAllister, R., Flege, J.E., Piske, T. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30, 229-258.
- [8] Pinet, M., Iverson, P. 2010. Talker-listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience. *J. Acoust. Soc. Am.* 128(3), 1357-1365.
- [9] van Wijngaarden, S.J., Steeneken, H.J.M., Houtgast, T. 2002. Quantifying the intelligibility of speech in noise for non-native talkers. *J. Acoust. Soc. Am.* 112(6), 3004-3013.
- [10] Zhang, Y., Nissen, S.L., Francis, A. 2008. Acoustics characteristics of English lexical stress produced by native Mandarin speakers. *J. Acoust. Soc. Am.* 123(6), 4498-4513.